Structural complexity notions for foundational theories

Dino Rossegger
Technische Universität Wien

Pennstate Logic Seminar

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101026834.

The complexity of a first-order theory can be measured in various ways.

- Quantifier complexity of axiomatization
- Turing degree
- Classification theory

- Complexity of the isomorphism relation
    - Borel reducibility
    - *Scott analysis*
- *Borel complexity of the set of models*

Scott analysis of models of arithmetic

Given a countable relational vocabulary τ, the set of countable τ-structures with universe ω admits a canonical Polish topology — the ***Vaught topology***.

Fix an enumeration $\varphi_i(x_0, \dots, x_i)$ of the atomic τ-formulas and let the atomic diagram of a τ-structure $\mathcal{A}$ with universe ω be

$$D(\mathcal{A})(i) = \begin{cases} 1 & \varphi_i[x_0 \dots x_i \mapsto 0 \dots i]^{\mathcal{A}} \\ 0 & \text{otherwise} \end{cases}$$

We get an homeomorphism $Mod(\tau) \to 2^\omega$ and can define the Borel hierarchy as usual:

For countable α, and $X \subseteq Mod(\tau)$

$$X \in \mathbf{\Sigma}^0_1 \iff X \text{ open} \quad X \in \mathbf{\Pi}^0_1 \iff X \text{ closed}$$

$$X \in \mathbf{\Sigma}^0_\alpha \iff X = \bigcup X_i \qquad , X_i \in \mathbf{\Pi}^0_{<\alpha}$$
$$X \in \mathbf{\Pi}^0_\alpha \iff X = \bigcap X_i \qquad , X_i \in \mathbf{\Sigma}^0_{<\alpha}$$

$L_{\omega_1\omega}$ is similar to first-order logic except it allows countable conjunctions and disjunctions.

For $\varphi \in L_{\omega_1\omega}$ and α countable

$$\varphi \in \Sigma_0^{\text{in}} = \Pi_0^{\text{in}} \iff \varphi \text{ finite and quantifier-free}$$
$$\varphi \in \Sigma_\alpha^{\text{in}} \iff \varphi = \bigvee\!\!\!\bigvee \exists \bar{x}_i \varphi_i \quad , \varphi_i \in \Pi_{<\alpha}^{\text{in}}$$
$$\varphi \in \Pi_\alpha^{\text{in}} \iff \varphi = \bigwedge\!\!\!\bigwedge \forall \bar{x}_i \varphi_i \quad , \varphi_i \in \Sigma_{<\alpha}^{\text{in}}$$

The ***asymmetric back-and-forth*** $\leq_\alpha$ relations are defined as

$$(\mathcal{A}, \bar{a}) \leq_1 (\mathcal{B}, \bar{b}) \iff \Pi_1\text{-tp}^{\mathcal{A}}(\bar{a}) \subseteq \Pi_1\text{-tp}^{\mathcal{B}}(\bar{b})$$
$$(\mathcal{A}, \bar{a}) \leq_\alpha (\mathcal{B}, \bar{b}) \iff (\forall \beta < \alpha) \forall \bar{c} \exists \bar{d}\, (\mathcal{B}, \bar{b}\bar{c}) \leq_\beta (\mathcal{A}, \bar{a}\bar{d})$$

Theorem (Karp 1965) $(A, \bar{a}) \leq_\alpha (\mathcal{B}, \bar{b})$ *if and only if* Π_α^{in}*-tp*$^{\mathcal{A}}(\bar{a}) \subseteq \Pi_\alpha^{\text{in}}$*-tp*$^{\mathcal{B}}(\bar{b})$.

Theorem (Scott 1964)
For every countable structure $\mathcal{A}$ there is an $L_{\omega_1\omega}$-sentence φ—the **Scott sentence** *of $\mathcal{A}$—such that for any countable $\mathcal{B}$, $\mathcal{B} \models \varphi$ if and only if $\mathcal{B} \cong \mathcal{A}$.*

Theorem (Montalbán 2015)
The following are equivalent for all $\alpha < \omega$ and countable $\mathcal{A}$:

1. *The isomorphism class of $\mathcal{A}$ is $\Pi^{\text{in}}_{\alpha+1}$.*
2. *There is a $\Pi^{\text{in}}_{\alpha+1}$ Scott sentence for $\mathcal{A}$.*
3. *The structure $\mathcal{A}$ is uniformly $\mathbf{\Delta}^0_\alpha$-categorical.*
4. *All automorphism orbits in $\mathcal{A}$ are $\Sigma^{\text{in}}_\alpha$ definable without parameters.*
5. *No tuple in $\mathcal{A}$ is α-free.*

The least α such that $\mathcal{A}$ satisfies any of the conditions is the **(parameterless) Scott rank** *of $\mathcal{A}$, $SR(\mathcal{A})$.*

A tuple $\overline{a}$ is **α-free** in $\mathcal{A}$ if $(\forall \beta < \alpha)\forall \overline{b} \exists \overline{a}', \overline{b}' \left(\overline{a}\overline{b} \leq_\beta \overline{a}'\overline{b}' \wedge \overline{a} \not\leq_\alpha \overline{a}' \right)$.

Definition (Makkai 1981)
The *Scott spectrum* of a theory T is the set

$$SSp(T) = \{\alpha \in \omega_1 : \text{there is a countable model of } T \text{ with Scott rank } \alpha\}.$$

Here T might be a sentence in $L_{\omega_1\omega}$.

Definition (Makkai 1981)
The *Scott spectrum* of a theory T is the set

$$SSp(T) = \{\alpha \in \omega_1 : \text{there is a countable model of } T \text{ with Scott rank } \alpha\}.$$

Here T might be a sentence in $L_{\omega_1\omega}$.

- Ash (1986) $SR(n) = 1$, $SR(\omega^\alpha) = 2\alpha$, $SR(\omega^\alpha + \omega^\alpha) = 2\alpha + 1$
  $\implies SSp(LO) = \omega_1 - \{0\}$
- The standard model $\mathbb{N}$ of PA has Scott rank 1: Every element is the nth successor of $\dot{0}$ for some $n \in \omega$, so the automorphism orbits are definable by $s(s(\dots(\dot{0})\dots)) = x$.
- $1 \in SSp(PA)$

What else can we say about the Scott spectra of PA and completions of PA?

Theorem (Montalbán, R. 2022)

1. $SSp(PA) = 1 \cup \{\alpha : \omega \leq \alpha < \omega_1\}$, $SSp(Th(\mathbb{N})) = 1 \cup \{\alpha : \omega < \alpha < \omega_1\}$, *and for* T *a non-standard completion of* PA, $SSp(T) = [\omega, \omega_1)$.
2. *If* $\mathcal{M}$ *is non-homogeneous, then* $SR(\mathcal{M}) \geq \omega + 1$.
3. *If* $\mathcal{M}$ *is non-standard atomic , then* $SR(\mathcal{M}) = \omega$.
4. *If* $\mathcal{M}$ *is non-standard homogeneous, then* $SR(\mathcal{M}) \in [\omega, \omega + 1]$.

Let $Tr_{\Delta_1^0}$ be a truth predicate for bounded formulas and define:

$$\bar{u} \leq_0^a \bar{v} \Leftrightarrow \forall(x \leq a)(Tr_{\Delta_1^0}(x, \bar{u}) \to Tr_{\Delta_1^0}(x, \bar{v}))$$

$$\bar{u} \leq_{n+1}^a \bar{v} \Leftrightarrow \forall \bar{x} \exists \bar{y} \Big(|\bar{x}| \leq a \to (|\bar{y}| \leq a \wedge \bar{u}\bar{x} \leq_n^a \bar{v}\bar{y}) \Big)$$

Let $Tr_{\Delta_1^0}$ be a truth predicate for bounded formulas and define:

$$\bar{u} \leq_0^a \bar{v} \Leftrightarrow \forall(x \leq a)(Tr_{\Delta_1^0}(x, \bar{u}) \rightarrow Tr_{\Delta_1^0}(x, \bar{v}))$$

$$\bar{u} \leq_{n+1}^a \bar{v} \Leftrightarrow \forall \bar{x} \exists \bar{y} \Big(|\bar{x}| \leq a \rightarrow (|\bar{y}| \leq a \wedge \bar{u}\bar{x} \leq_n^a \bar{v}\bar{y}) \Big)$$

Proposition

Let $\bar{a}, \bar{b} \in M$. Then $\bar{a} \leq_n \bar{b} \Leftrightarrow \forall(m \in \omega)\mathcal{M} \models \bar{a} \leq_n^{\dot{m}} \bar{b}$. Furthermore, if there is $c \in M - \mathbb{N}$ such that $\mathcal{M} \models \bar{a} \leq_n^c \bar{b}$, then $\bar{a} \leq_n \bar{b}$.

Lemma *For every $\bar{a}, \bar{b} \in M^{<\omega}$, $\bar{a} \leq_\omega \bar{b}$ if and only if $tp(\bar{a}) = tp(\bar{b})$.*

Proof sketch.
($\Rightarrow$) The conjunction over all formulas in a type is Π_ω^{in} and $\bar{a} \leq_\omega \bar{b}$ iff Π_ω^{in}-tp$(\bar{a}) \subseteq \Pi_\omega^{\text{in}}$-tp$(\bar{b})$.
($\Leftarrow$) Take $\mathcal{N} \succeq \mathcal{M}$ homogeneous, then $tp(\bar{a}) = tp(\bar{b}) \implies \bar{b} \in aut_{\mathcal{N}}(\bar{a})$, so $\bar{a} \leq_\omega \bar{b}$ and for all n, m, $\bar{a} \leq_n^{\dot{m}} \bar{b}$. This also holds in $\mathcal{M}$. Lemma follows by definition of $\leq_\omega$. □

The lemma implies that non-homogeneous models of PA cannot have Scott rank $\leq \omega$ as they contain $\overline{a}, \overline{b}$ with $tp(\overline{a}) = tp(\overline{b})$, hence $\overline{a} \leq_\omega \overline{b}$ and $\overline{a} \notin aut(\overline{b})$.

Using the definability of the formal back-and-forth relations and elementary extending to non–homogeneous models we get the lower bounds.

Lemma
If $\mathcal{M}$ is a non-standard model of PA then $SR(\mathcal{M}) \geq \omega$.

If for any countable well-order L we can find a model $\mathcal{M}_L$ such that L is ***infinitary bi-interpretable*** with $\mathcal{M}_L$ by formulas of the right complexity then we would get $SR(\mathcal{M}_L)) = \omega + SR(L)$.

Theorem (Gaifman 1976)
Let T be a completion of PA and L a linear order. Then there is a model $\mathcal{M}_L$ of T with $Aut(\mathcal{N}_L) \cong Aut(L)$.

Theorem (Harrison-Trainor, Montalbán, Miller 2018)
Two countable structures $\mathcal{A}$ and $\mathcal{B}$ are infinitary bi-interpretable if and only if their automorphism groups are isomorphic.

A careful analysis of Gaifman's theorem shows that the complexity of the formulas involved in the bi-interpretation is just right so that we can find models $\mathcal{M}_L$ with $SR(\mathcal{M}_L) = \omega + SR(L)$.

Theorem (Montalbán, R. 2022)

1. $SSp(PA) = 1 \cup \{\alpha : \omega \leq \alpha < \omega_1\}$, $SSp(Th(\mathbb{N})) = 1 \cup \{\alpha : \omega < \omega < \omega_1\}$, *and for* T *a non-standard completion of* PA, $SSp(T) = [\omega, \omega_1)$.
2. *If* $\mathcal{M}$ *is non-homogeneous, then* $SR(\mathcal{M}) \geq \omega + 1$.
3. *If* $\mathcal{M}$ *is non-standard atomic , then* $SR(\mathcal{M}) = \omega$.
4. *If* $\mathcal{M}$ *is non-standard homogeneous, then* $SR(\mathcal{M}) \in [\omega, \omega + 1]$.

Theorem (Montalbán, R. 2022)

1. $SSp(PA) = 1 \cup \{\alpha : \omega \leq \alpha < \omega_1\}$, $SSp(Th(\mathbb{N})) = 1 \cup \{\alpha : \omega < \omega < \omega_1\}$, *and for* T *a non-standard completion of* PA, $SSp(T) = [\omega, \omega_1)$.
2. *If* $\mathcal{M}$ *is non-homogeneous, then* $SR(\mathcal{M}) \geq \omega + 1$.
3. *If* $\mathcal{M}$ *is non-standard atomic , then* $SR(\mathcal{M}) = \omega$.
4. *If* $\mathcal{M}$ *is non-standard homogeneous, then* $SR(\mathcal{M}) \in [\omega, \omega + 1]$.

- SSp for completions of $PA^{-} + I\Sigma_n$?
- SSp for other foundational theories $Z_2, KP, ZF, ...$?
- Are there non-atomic homogeneous models of PA of Scott rank ω?

Theorem (Łełyk, Szlufik 2023; in preparation)
If $\mathcal{M}$ is a homogeneous model of PA that is not atomic, then $SR(\mathcal{M}) = \omega + 1$.

Kalociński 2023: Many theories have an ***intended model***, a model that we have in mind when axiomatizing the theory. In the case of PA, this intendedness can be seen in the Scott analysis. Does this phenomenon also appear in other theories? Can we discover intended models from Scott analysis?

Theorem (Łełyk, Szlufik 2023; in preparation)
If $\mathcal{M}$ is a homogeneous model of PA that is not atomic, then $SR(\mathcal{M}) = \omega + 1$.

Kalociński 2023: Many theories have an ***intended model***, a model that we have in mind when axiomatizing the theory. In the case of PA, this intendedness can be seen in the Scott analysis. Does this phenomenon also appear in other theories? Can we discover intended models from Scott analysis?

Already for fragments of second-order arithmetic, one can see that Scott analysis does not reveal intendedness.

For RCA_0 ($PA^- + I\Sigma^0_1 + \Delta^0_1$-$CA$), $(\mathbb{N}, S) \models RCA_0$ when S is a countable Turing ideal and all of these models have same Scott rank.

$\implies$ *computable Scott rank*

The complexity of the set of models of a theory

So far we have attempted to measure the complexity of T via its isomorphism relation. Why not look at the descriptive complexity of its set of models?

Definition
Let X be a Polish space and $A \subseteq X$, then for any point class Γ, A is *Γ-complete* if $A \in \Gamma(X)$ and for every $B \in \Gamma(Y)$ for any Polish Y, B is *Wadge reducible* to A, $B \leq_W A$, i.e., there is continuous $f : Y \to X$ with $f(y) \in X$ if and only $y \in Y$.

For any theory T, $Mod(T) \in \mathbf{\Pi}^0_\omega$.
Is $Mod(TA)$, the set of models of true arithmetic $\mathbf{\Pi}^0_\omega$-complete?

Theorem (Andrews, Lempp, R. in preparation)
For any complete first-order theory T, $Mod(T)$ is $\mathbf{\Pi}^0_\omega$ complete if and only if T has no axiomatization by formulas of bounded quantifier-complexity.

The theorem suggests that $L_{\omega_1\omega}$ is not more efficient when talking about sets of models. I.e.,

Corollary
If T is not bounded axiomatizable, then $Mod(T)$ is is not Σ^{in}_n definable for any $n \in \omega$.

Proof.
Assume it was, then by Lopez-Escobar $Mod(T)$ is $\mathbf{\Sigma}^0_n$ and thus not $\mathbf{\Pi}^0_\omega$-complete. So it is axiomatizable by formulas of bounded quantifier-complexity.

□

($\Leftarrow$) Say, S is a set of Σ_n formulas axiomatizing $Mod(T)$, then $\bigwedge_{\varphi \in S} \varphi$ is Π^{in}_{n+1} and hence by Lopez-Escobar, $Mod(T)$ is $\mathbf{\Pi}^0_{n+1}$.
($\Rightarrow$) This direction relies on an old theorem due to Solovay.

Theorem (Solovay 1982)
Let T be a complete theory. Suppose $R \leq_T X$ is an enumeration of a Scott set S, with functions t_n which are $\Delta^0_n(X)$ uniformly in n, such that for each n, $\lim_s t_n(s)$ is an R-index for $T \cap \Sigma_n$, and for all s, $t_n(s)$ is an R-index for a subset of $T \cap \Sigma_n$. Then T has a model $\mathcal{B}$, representing S, with $\mathcal{B} \leq_T X$.

- Known proofs use methods for iterated Priority constructions
- Original proof uses a Harrington style worker argument
- Version above is due to Knight (1999) and proved using version of α-systems

A *Scott set* $S \subseteq 2^\omega$ is a set satisfying

1. $x \leq_T y$ and $y \in S \implies x \in S$,
2. $x, y \in S \implies x \oplus y \in S$,
3. and if $x \in S$ codes an infinite binary tree T_x, then $S \cap [T_x] \neq \emptyset$.

A *Scott set* $S \subseteq 2^\omega$ is a set satisfying

1. $x \leq_T y$ and $y \in S \implies x \in S$,
2. $x, y \in S \implies x \oplus y \in S$,
3. and if $x \in S$ codes an infinite binary tree T_x, then $S \cap [T_x] \neq \emptyset$.

$R \in 2^\omega$ is an *enumeration* of a Scott set S if $\{R^{[i]} : i \in \omega\} = S$.

A ***Scott set*** $S \subseteq 2^\omega$ is a set satisfying

1. $x \leq_T y$ and $y \in S \implies x \in S$,
2. $x, y \in S \implies x \oplus y \in S$,
3. and if $x \in S$ codes an infinite binary tree T_x, then $S \cap [T_x] \neq \emptyset$.

$R \in 2^\omega$ is an ***enumeration*** of a Scott set S if $\{R^{[i]} : i \in \omega\} = S$.

A countable model $\mathcal{M}$ ***represents*** a countable Scott set S if for all complete *B_n-types* $\Gamma(\bar{u}, x)$ and all $\bar{c} \in M$:

$$\Gamma(\bar{c}, x) \text{ realized in } \mathcal{M} \iff \Gamma \in S \text{ and } Con(\Gamma(\bar{c}, x) \cup Diag_{el}(\mathcal{M})).$$

Fix a theory T not axiomatizable by bounded quantifier formulas and theories $T_n \neq T$ such that $T_n \cap \Sigma_n = T \cap \Sigma_n$, an enumeration R of a Scott set S containing T, (T_n) and a Borel code C for a fixed $\mathbf{\Pi}^0_\omega$ set $P = \bigcap P_n$ where P_n is Σ_n.

In order to prove our theorem we:

- Given x produce (indices) for functions t_n such that $t_n(x^{(n-1)}, s) = R(T_{n+1})$ if $x \notin P_{n,s}$ and $t_n(x^{(n-1)}, s) = R(T)$ otherwise. This can be done recursive in $x \oplus S \oplus C$.
- Verify that Solovay's theorem is continuous, i.e., the function $T \mapsto \mathcal{B}$ is continuous.

Corollary
$Mod(PA)$, and $Mod(T)$ for T a completion of PA are $\mathbf{\Pi}^0_\omega$-complete.

Proof.
By Tarski's undefinability of truth, no completion of PA is axiomatizable by formulas of bounded quantifier-complexity. To get $Mod(PA)$ take $T = TA$ and let T_n such that $T_n \nvDash I\Sigma_n$.

□

What about other theories, what role does induction play in this? How about completions of PA^-? We asked Roman Kossak who asked Ali Enayat and Albert Visser.

What about other theories, what role does induction play in this? How about completions of PA^-? We asked Roman Kossak who asked Ali Enayat and Albert Visser.

Definition (Pudlák 1983, Pakhomov and Visser 2022)
A (possibly incomplete) τ-theory T is ***sequential*** if it admits a definitional extension to ***Adjunctive set theory*** $\mathrm{AS}(T)$, namely, in $\tau \sqcup \{\in\}$, we have the axioms

1. $\exists x \, \forall y \, (\neg y \in x)$ ("the empty set exists"), and
2. $\forall x \, \forall y \, \exists z \, \forall w \, (w \in z \leftrightarrow (w \in x \vee w = y))$ ("$x \cup \{y\}$ exists").

What about other theories, what role does induction play in this? How about completions of PA^-? We asked Roman Kossak who asked Ali Enayat and Albert Visser.

Definition (Pudlák 1983, Pakhomov and Visser 2022)
A (possibly incomplete) τ-theory T is ***sequential*** if it admits a definitional extension to ***Adjunctive set theory*** $\mathrm{AS}(T)$, namely, in $\tau \sqcup \{\in\}$, we have the axioms

1. $\exists x \, \forall y \, (\neg y \in x)$ ("the empty set exists"), and
2. $\forall x \, \forall y \, \exists z \, \forall w \, (w \in z \leftrightarrow (w \in x \vee w = y))$ ("$x \cup \{y\}$ exists").

In essence, sequential theories allow for coding of finite sequences as in Gödel's β-function (but do not require, e.g., extensionality).

What about other theories, what role does induction play in this? How about completions of PA^-? We asked Roman Kossak who asked Ali Enayat and Albert Visser.

Definition (Pudlák 1983, Pakhomov and Visser 2022)
A (possibly incomplete) τ-theory T is ***sequential*** if it admits a definitional extension to ***Adjunctive set theory*** $\mathrm{AS}(T)$, namely, in $\tau \sqcup \{\in\}$, we have the axioms

1. $\exists x \, \forall y \, (\neg y \in x)$ ("the empty set exists"), and
2. $\forall x \, \forall y \, \exists z \, \forall w \, (w \in z \leftrightarrow (w \in x \vee w = y))$ ("$x \cup \{y\}$ exists").

In essence, sequential theories allow for coding of finite sequences as in Gödel's β-function (but do not require, e.g., extensionality).

Examples of sequential theories:

PA, $I\Delta_0 + \exp$, ZF, KP, *even* PA^- (Jeřábek 2012), $\mathrm{AS} = \mathrm{AS}(\emptyset)$ (Pakhomov, Visser 2022), but *not* Robinson's Q.

Theorem (Enayat, Visser in preparation)
No sequential theory in finite vocabulary has an axiomatization by sentences of bounded quantifier complexity.

The finiteness condition here is essential. Consider the Morleyization of true arithmetic (add a relation R_φ for every formula φ). This has a compositional axiomatization in the style of Tarski's definition of satisfaction, and hence an axiomatization by Π_2 formulas.

Corollary
If T is sequential in finite vocabulary, then $Mod(T)$ is $\mathbf{\Pi}^0_\omega$ complete.

Our techniques can be used to give an alternative proof to a remarkable result by Harrison-Trainor and Kretschmer which is another witness that $L_{\omega_1\omega}$ is not more efficient than first-order logic.

Corollary (Harrison-Trainor, Kretschmer 2022; ALR, Gonzalez, Zhu in preparation)
If φ is a first-order formula that is equivalent to a Σ_n^{in} formula, then φ is equivalent to a Σ_n formula.

We are currently working on extending this to a full characterization of the complexity of a theories' set of models in terms of first-order axiomatization.

Our techniques can be used to give an alternative proof to a remarkable result by Harrison-Trainor and Kretschmer which is another witness that $L_{\omega_1\omega}$ is not more efficient than first-order logic.

Corollary (Harrison-Trainor, Kretschmer 2022; ALR, Gonzalez, Zhu in preparation)
If φ is a first-order formula that is equivalent to a Σ_n^{in} formula, then φ is equivalent to a Σ_n formula.

We are currently working on extending this to a full characterization of the complexity of a theories' set of models in terms of first-order axiomatization.

Thank you!